Kapitel 10

Statistik

10.1 Wahrscheinlichkeit

Das Ergebnis einer Messung oder Beobachtung wird Ereignis genannt. Ereignisse werden mit den Buchstaben A, $B, \ldots$ bezeichnet. Die Messung einer kontinuierlichen Variablen x gibt in der Regel (nicht abzählbar) unendlich viele verschiedene Ereignisse, die jedoch zu abzählbar vielen Ereignissen zusammen gefaßt werden können (für eine kontinuierliche Zufallsvariable x kann z.B. das Ereignis $0.5 \leq x \leq 1.0$ betrachtet werden). Das Ereignis, daß die Messung irgendein Ereignis liefert, wird als Einheitsereignis E bezeichnet.

Eine plausible Definition der Wahrscheinlichkeit $P(A)$ für das Auftreten von Ereignis A ist der Quotient aus der Zahl n von Beobachtungen von A und der Gesamtheit N der Beobachtungen im Limes $N \to \infty$:

$$P(A) = \lim_{N\to\infty} \frac{n}{N}$$

Folgende Schreibweisen werden benutzt:

Ereignis	Wahrscheinlichkeit
A	$P(A)$
A oder B	$P(A+B)$
A und B	$P(AB)$
B, wenn A	$P(B\|A)$

Der Wahrscheinlichkeitstheorie können die folgenden Axiome zugrunde gelegt werden:

- $P(A) \geq 0$
- $P(E) = 1$
- $P(A+B) = P(A) + P(B)$, wenn die Ereignisse A und B sich gegenseitig ausschließen.
- $P(AB) = P(A) \cdot P(B|A)$

Aus den Axiomen folgen die Aussagen:
Für komplementäre Ereignisse A und $\bar{A}$ gilt:

$$P(A+\bar{A}) = P(A) + P(\bar{A}) = 1 \qquad 0 \leq P(A) \leq 1$$

Zwei Ereignisse A und B heißen unabhängig, wenn $P(B)$ nicht davon abhängt, ob A bei der gleichen Beobachtung eingetreten ist oder nicht:

$$P(B|A) = P(B) \qquad P(AB) = P(A) \cdot P(B)$$

10.2 Zufallsvariable

Meßgrößen sind Zufallsvariablen. Alle Messungen oder Beobachtungen unterliegen zufälligen Schwankungen, daher sind die Zahlenwerte von Messungen oder Beobachtungen nicht exakt vorhersagbar; Grund ist die begrenzte Meßgenauigkeit oder der statistische Charakter der untersuchten Größe selbst. Je nachdem die Meßgrößen x

kontinuierliche oder diskrete Werte annehmen können, werden sie kontinuierliche oder diskrete Zufallsvariablen genannt. Beiden liegt eine statistische Verteilung zugrunde.

Die Verteilung einer *kontinuierlichen* Zufallsvariablen x wird bestimmt durch ihre Dichtefunktion $f(x)$ mit den Eigenschaften

$$f(x) \geq 0 \qquad \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Die durch

$$F(x) = \int_{-\infty}^{x} f(x')\,dx'$$

definierte Funktion $F(x)$ heißt Verteilungsfunktion der Zufallsvariablen x. Es gilt:

$$f(x) = \frac{dF(x)}{dx}$$

Die Wahrscheinlichkeit, daß ein Einzelwert x in das Intervall $[x_1, x_2]$ fällt, ist gegeben durch

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x)\,dx = F(x_2) - F(x_1)$$

Wichtige Eigenschaften einer Verteilung lassen sich durch wenige Parameter angeben. Die beiden wichtigsten Parameter einer Verteilung sind

- Mittelwert (Positionsparameter), und
- Standardabweichung (Streuungsparameter).

Der Mittelwert oder Erwartungswert der Größe x wird mit μ bezeichnet und ist definiert durch ($f(x) =$ Wahrscheinlichkeitsdichte)

$$\mu = E[x] = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx.$$

Allgemein sind Erwartungswerte von Funktionen $g(x)$ der Zufallsvariablen x definiert durch

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) \cdot f(x)\,dx.$$

Der Mittelwert μ ist der Erwartungswert für $g(x) = x$. Die Standardabweichung, bezeichnet mit σ, ist die Quadratwurzel aus der Varianz, die definiert ist als der Erwartungswert von $g(x) = (x - \mu)^2$:

$$\sigma^2 = V(x) = E[(x-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)\,dx.$$

Die Varianz σ^2 der Verteilung einer Zufallsvariablen x kann durch die Erwartungswerte von x und x^2 ausgedrückt werden:

$$\begin{aligned} \sigma^2 &= E[(x-\mu)^2] = E[x^2 - 2x\mu + \mu^2] = \int_{-\infty}^{+\infty} x^2 \cdot f(x)\,dx - 2\mu \int_{-\infty}^{+\infty} x \cdot f(x)\,dx + \mu^2 \int_{-\infty}^{+\infty} f(x)\,dx \\ &= E[x^2] - 2\mu^2 + \mu^2 = E[x^2] - (E[x])^2\,. \end{aligned}$$

Die gemeinsame Verteilung von zwei Zufallsvariablen x und y wird durch die Dichtefunktion $f(x, y)$ mit der Normierung

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$$

bestimmt. Die Mittelwerte μ_x und μ_y der beiden Variablen x und y sind definiert durch

$$\mu_x = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot f(x,y)\,dx\,dy \qquad \mu_y = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \cdot f(x,y)\,dx\,dy$$

Neben den Varianzen σ_x^2 und σ_y^2, definiert durch

$$\sigma_x^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\mu_x)^2 \cdot f(x,y)\,dx\,dy \qquad \sigma_y^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y-\mu_y)^2 \cdot f(x,y)\,dx\,dy$$

gibt es noch die Kovarianz σ_{xy}, definiert durch

$$\sigma_{xy} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\mu_x)(y-\mu_y) \cdot f(x,y)\,dx\,dy$$

Der durch $\rho = \sigma_{xy}/(\sigma_x \sigma_y)$ definierte Korrelationskoeffizient kann Werte zwischen +1 und -1 annehmen. Bei $\rho = 0$ heißen die beiden Variablen x und y unkorreliert. Wenn die beiden Variablen statistisch unabhängig sind, läßt sich ihre Wahrscheinlichkeitsdichte in der Form

$$f(x,y) = f_1(x) \cdot f_2(y)$$

schreiben und der Korrelationskoeffizient ist 0.

Bei einer *diskreten* Zufallsvariablen sind nur diskrete Werte $x_i, i = 1, 2\ldots$ möglich. Jedem möglichen Wert x_i kann eine Wahrscheinlichkeit $P(i)$ zugeordnet werden mit den Eigenschaften:

$$P(i) \geq 0 \qquad \sum_i P(i) = 1$$

Mittelwert und Varianz sind durch Summen über alle möglichen Werte definiert:

$$\begin{aligned} \mu &= E[x] = \sum_i x_i \cdot P(i) \\ \sigma^2 &= V(x) = \sum_i (x_i - \mu)^2 \cdot P(i) = E[x^2] - (E[x])^2 \,. \end{aligned}$$

Theoretische Verteilungen

Normalverteilung. Die Normalverteilung, auch Gaußverteilung genannt, wird durch die beiden Parameter Mittelwert μ und Standardabweichung σ vollständig festgelegt; die Dichtefunktion lautet:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

Aus dieser Dichte ergeben sich durch Integration die folgenden Aussagen über die Abweichung eines Einzelwertes x vom Mittelwert μ um ein, zwei und drei Standardabweichungen:

$$\begin{aligned} P(\mu - 1\sigma \leq x \leq \mu + 1\sigma) &= 68.27\% \\ P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) &= 95.45\% \\ P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) &= 99.73\% \end{aligned}$$

Bei der Normalverteilung ist die volle Breite bei halbem Maximalwert (Halbwertsbreite) gleich $2.34\,\sigma$. Die Verteilungsfunktion der Normalverteilung ist

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \frac{1}{2} + \frac{1}{2} erf\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$$

Viele Verteilungen der Praxis kommen der Normalverteilung sehr nahe und werden daher durch die Normalverteilung approximiert.

Gleichverteilung. Die Dichte der Gleichverteilung zwischen den Grenzen a und b ist

$$f(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

Für Mittelwert und Standardabweichung ergeben sich die Werte:

$$\mu = \frac{a+b}{2} \qquad \sigma = \frac{b-a}{\sqrt{12}}$$

Exponentialverteilung. Die Dichte der Exponentialverteilung hat nur einen Parameter λ:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Für Mittelwert und Standardabweichung ergeben sich:

$$\mu = \frac{1}{\lambda} \qquad \sigma = \frac{1}{\lambda}$$

Die Exponentialverteilung beschreibt z. B. die Häufigkeit von Ereignissen, die zeitlich zufällig mit einer konstanten Wahrscheinlichkeit erfolgen.

Binomialverteilung. Ein Ereignis A trete bei einem Versuch mit der Wahrscheinlichkeit p auf; entsprechend ist die Wahrscheinlichkeit für das Nicht-Auftreten von A gleich $q = 1 - p$. Gesucht ist die Wahrscheinlichkeit, daß bei n Versuchen das Ereignis A k mal auftritt. Die Wahrscheinlichkeit, daß in den ersten k Versuchen das Ereignis A und in den restlichen $n - k$ Versuchen das Ereignis $\bar{A}$ auftritt, ist das Produkt $p^k q^{n-k}$. Das Ereignis k mal A bei n Versuchen kann in $\binom{n}{k}$ verschiedenen Reihenfolgen auftreten. Daher ist die Wahrscheinlichkeit, daß bei n Versuchen k mal das Ereignis A auftritt, gegeben durch

$$P(k) = \binom{n}{k} p^k q^{n-k}$$

Dies entspricht dem Binomischen Lehrsatz

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = 1 \qquad\qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Mittelwert und Varianz der Binomialverteilung sind

$$E[k] = \mu = np \qquad \sigma^2 = npq$$

Poissonverteilung. Wenn man in der Binomialverteilung $n \to \infty$ wachsen läßt, dabei jedoch den Mittelwert $\mu = np$ konstant hält, geht die Binomialverteilung in die Poissonverteilung über:

$$P(k) = \frac{\mu^k}{k!} e^{-\mu}$$

Mittelwert und Varianz sind

$$E[k] = \mu \qquad \sigma^2 = \mu \qquad \sigma = \sqrt{\mu}$$

Die (diskrete) Poissonverteilung geht für große Werte von μ über in die spezielle Normalverteilung ($\sigma = \sqrt{\mu}$)

$$P(k) = \frac{1}{\sqrt{2\pi\mu}} e^{-(k-\mu)^2/2\mu}$$

10.3 Funktionen von Zufallsvariablen

Funktionen von Zufallsvariablen sind selbst wieder Zufallsvariable, für die Mittelwerte und Varianzen angegeben werden können. Die Berechnung der Varianz einer Funktion von Zufallsvariablen aus den Varianzen der Veränderlichen der Funktion nennt man Fehlerfortpflanzung (error propagation).

Funktionen einer Zufallsvariablen. Betrachtet wird eine Funktion $w = w(x)$ der Zufallsvariablen x, die einer Verteilung mit Mittelwert μ_x und Standardabweichung σ_x folgt. Der Mittelwert μ_w kann wie folgt berechnet werden. Ausgehend von der Taylor-Entwicklung

$$w(x) \approx w(\mu_x) + \left.\frac{dw}{dx}\right|_{\mu_x} (x - \mu_x)$$

ergibt sich als Mittelwert von w:

$$\mu_w = E[w] \approx w(\mu_x) + \left.\frac{dw}{dx}\right|_{\mu_x} E[x - \mu_x] \quad = w(\mu_x)$$

Für lineare Funktionen ist diese Beziehung exakt; bei nichtlinearen Funktionen $w(x)$ würde zwar der quadratische Term der Taylorentwicklung eine (kleine) Korrektur liefern, diese ist in der Praxis jedoch i.a. vernachlässigbar. Für die Varianz von w ergibt sich bei Benutzung der Taylorentwicklung als Erwartungwert von $(w(x) - w(\mu_x))^2$:

$$\sigma_w^2 = E\left[(w(x) - \mu_w)^2\right] = E\left[(x - \mu_x)^2\right] \left(\left. \frac{dw}{dx} \right|_{\mu_x} \right)^2 = \sigma_x^2 \left(\left. \frac{dw}{dx} \right|_{\mu_x} \right)^2 .$$

Daher folgt:

$$\sigma_w \approx \sigma_x \left| \frac{dw}{dx} \right|_{x=\mu_x}$$

Funktionen mehrerer Zufallsvariabler. Betrachtet wird zunächst eine Funktion $w(x, y)$ von zwei Zufallsvariablen x und y, die Mittelwerte μ_x und μ_y und Standardabweichungen σ_x und σ_y haben. Entsprechend der linearen Näherung

$$w(x, y) \approx w(\mu_x, \mu_y) + \left. \frac{\partial w}{\partial x} \right|_{\mu_x, \mu_y} (x - \mu_x) \quad + \left. \frac{\partial w}{\partial y} \right|_{\mu_x, \mu_y} (y - \mu_y)$$

ergibt sich für Mittelwert und Varianz der Größe w:

$$\begin{aligned} \mu_w &\approx w(\mu_x, \mu_y) \\ \sigma_w^2 &\approx \sigma_x^2 \left(\left. \frac{\partial w}{\partial x} \right|_{\mu_x, \mu_y} \right)^2 + \sigma_y^2 \left(\left. \frac{\partial w}{\partial y} \right|_{\mu_x, \mu_y} \right)^2 \end{aligned}$$

Diese Formeln gelten für den Fall von *unabhängigen* Zufallsvariablen x und y. Wenn die Zufallsvariablen x und y *nicht* statistisch unabhängig sind, ergibt sich für den Ausdruck von σ_w^2 ein (positiver oder negativer) Zusatzterm:

$$\sigma_w^2 \approx \sigma_x^2 \left(\left. \frac{\partial w}{\partial x} \right|_{\mu_x, \mu_y} \right)^2 + 2\sigma_{xy} \left(\left. \frac{\partial w}{\partial x} \right|_{\mu_x, \mu_y} \right) \cdot \left(\left. \frac{\partial w}{\partial y} \right|_{\mu_x, \mu_y} \right) + \sigma_y^2 \left(\left. \frac{\partial w}{\partial y} \right|_{\mu_x, \mu_y} \right)^2$$

Die Verallgemeinerung auf eine Funktion w von n statistisch unabhängigen Zufallsvariablen x_1, $x_2 \ldots x_n$ ergibt:

$$\begin{aligned} \mu_w &\approx w(\mu_1, \mu_2 \ldots \mu_n) \\ \sigma_w^2 &\approx \sum_i \sigma_i^2 \left(\left. \frac{\partial w}{\partial x_i} \right|_{\mu_1, \ldots \mu_n} \right)^2 \end{aligned}$$

Spezialfälle. Wenn w von der Form

$$w = x + y - z \cdots$$

(Summen und Differenzen) ist, ergibt die Formel

$$\sigma_w^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \cdots$$

wenn σ_x, σ_y, $\sigma_z \ldots$ die Standardabweichungen der unabhängigen Zufallsvariablen x, y, $z \ldots$ sind; die Quadrate der einzelnen Standardabweichungen addieren sich.
Wenn w von der Form

$$w = \frac{x \cdot y \cdots}{z \cdots}$$

(Produkte und Quotienten) ist, ergibt die Formel bei unabhängigen Zufallsvariablen nach Umformung den Ausdruck

$$\left(\frac{\sigma_w}{\mu_w} \right)^2 = \left(\frac{\sigma_x}{\mu_x} \right)^2 + \left(\frac{\sigma_y}{\mu_y} \right)^2 + \left(\frac{\sigma_z}{\mu_z} \right)^2 + \cdots$$

Die Quadrate der *relativen* Standardabweichungen addieren sich.
Für die Funktion (Mittelwert)

$$w = \frac{1}{n} \left(x_1 + x_2 + \ldots + x_n \right),$$

wobei alle Zufallsvariablen x_i der gleichen Verteilung mit Mittelwert μ_x und Standardabweichung σ_x entstammen und statistisch unabhängig sind, gilt:

$$\sigma_w^2 = \sum_{i=1}^{n} \sigma_x^2 \frac{1}{n^2} = \frac{\sigma_x^2}{n}.$$

Die Standardabweichung des Mittelwerts von n unabhängigen Zufallsvariablen nimmt also bei Vergrößerung der Zahl n der Messungen proportional zu $1/\sqrt{n}$ ab.
Zentraler Grenzwertsatz: Sind die Zufallsgrößen x_i unabhängig verteilt mit Mittelwert μ und Varianz σ_x^2, so ist der Mittelwert

$$\frac{1}{n} \sum_{i=1}^{n} x_i$$

im Grenzfall $n \to \infty$ *normalverteilt* mit Mittelwert μ und Varianz σ_x^2/n.

10.4 Auswertung von Messungen

Ein Meßwert x einer Meßgröße mit wahrem Wert μ folge einer Normalverteilung mit der durch die Meßapparatur bedingten Standardabweichung σ (Meßfehler). Ausgehend von der für die Normalverteilung geltenden Aussage

$$P(\mu - \sigma \leq x \leq \mu + \sigma) = 68.27\%$$

erhält man durch Umformung der Ungleichungen

$$\mu - \sigma \leq x \qquad x \leq \mu + \sigma$$

in die Ungleichungen

$$\mu \leq x + \sigma \qquad x - \sigma \leq \mu$$

die Aussage

$$P(x - \sigma \leq \mu \leq x + \sigma) = 68.27\%$$

Diese Aussage bedeutet, daß der wahre Wert μ mit einer Wahrscheinlichkeit von 68.27 % innerhalb der (einfachen) durch die Standardabweichung gegebenen Fehlergrenzen um den Meßwert liegen. Die Annahme einer Normalverteilung ist bei Messungen meist gerechtfertigt (eine theoretische Begründung liefert der zentrale Grenzwertsatz). Meßergebnisse werden in der Form *Meßwert* $\pm$ *Fehler*

$$x \pm \sigma$$

angegeben, wobei als Fehler die (einfache) Standardabweichung angegeben wird.
Empirische Werte von Mittelwert und Varianz. Zur Bestimmung von Parametern der zugrunde liegenden Verteilung aus statistisch verteilten Daten x_i, $i = 1, \ldots n$ werden Schätzfunktionen $t(x_1, x_2, \ldots x_n)$ benutzt; diese sind als Funktionen der Daten selbst Zufallsvariablen und sollten u.a. die Eigenschaft haben, daß ihre Erwartungswerte im Grenzwert $n \to \infty$ (Konsistenz) und bei endlichen Werten von n (Erwartungstreue) gleich den zu schätzenden Parametern der Verteilung ist. Eine Schätzfunktion m für den Mittelwert einer Meßreihe x_i, $i = 1, \ldots n$ ist

$$m = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Der Erwartungswert von m ist

$$E[m] = \frac{1}{n} \sum_{i=1}^{n} E[x_i] = \mu$$

bei $E[x_i] = \mu$ und damit ist m eine erwartungstreue Schätzfunktion. Eine erwartungstreue Schätzfunktion für die Varianz ist

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$$

Zum Nachweis der Erwartungstreue wird zunächst die Summe umgeformt:

$$\begin{aligned}\sum_{i=1}^{n}(x_i-m)^2 &= \sum_{i=1}^{n}\left((x_i-\mu)-(m-\mu)\right)^2 = \sum_{i=1}^{n}(x_i-\mu)^2 - 2(m-\mu)\sum_{i=1}^{n}(x_i-\mu)+n(m-\mu)^2 \\ &= \sum_{i=1}^{n}(x_i-\mu)^2 - n(m-\mu)^2 \ ,\end{aligned}$$

denn es gilt

$$\sum_{i=1}^{n}(x_i-\mu) = n(m-\mu).$$

Der Erwartungswert der Summe ist:

$$E\left[\sum_{i=1}^{n}(x_i-m)^2\right] = E\left[\sum_{i=1}^{n}(x_i-\mu)^2\right] - n\cdot E\left[(m-\mu)^2\right] = n\sigma_x^2 - n\frac{\sigma_x^2}{n} = (n-1)\sigma_x^2.$$

Also ist der Erwartungswert von s_x^2

$$E[s_x^2] = \frac{1}{n-1}(n-1)\sigma_x^2 = \sigma_x^2$$

und damit ist s_x^2 eine erwartungstreue Schätzfunktion für σ_x^2. Die empirische Standardabweichung s_x wird oft wie die wahre Standardabweichung behandelt. Tatsächlich ist sie jedoch als Funktion von Zufallsvariablen selbst eine Zufallsvariable und damit statistischen Schwankungen unterworfen. Für große Werte von n (etwa $n \geq 10$) ist die statistische Schwankung gering und meist vernachlässigbar, bei sehr kleinen Werten von n ist jedoch Vorsicht geboten.

Formeln. Die folgenden Formeln sind anwendbar zur Berechnung von Mittelwert und Standardabweichungen, wenn n Einzelwerte $x_1,\ x_2 \ldots x_n$ gleicher Genauigkeit vorliegen. Die Formeln enthalten zur Erhöhung der numerischen Genauigkeit einen geeignet zu wählenden Wert x_0, der ungefähr gleich dem Mittelwert sein sollte. Gebildet werden zunächst die Summen:

$$S_x = \sum_{i=1}^{n}(x_i-x_0) \qquad S_{xx} = \sum_{i=1}^{n}(x_i-x_0)^2$$

Der Mittelwert $\bar{x}$, die Standardabweichung des Mittelwerts $s_{\bar{x}}$ und die Standardabweichung der Einzelwerte s_x ergeben sich aus:

$$\bar{x} = x_0 + \frac{S_x}{n} \qquad s_{\bar{x}}^2 = \frac{1}{n(n-1)}\left(S_{xx} - \frac{S_x^2}{n}\right) \qquad s_x^2 = \frac{1}{n-1}\left(S_{xx} - \frac{S_x^2}{n}\right)$$

Geradenanpassung. Bei Meßreihen wird häufig eine Zufallsvariable y als Funktion von einer jeweils fest einstellbaren Größe x gemessen. Wenn für die Größe y eine funktionelle Abhängigkeit von x bekannt ist, können die Daten $x_i,\ y_i,\ i = 1,\ 2\ldots n$ benutzt werden, um Parameter, die in der funktionellen Abhängigkeit vorkommen, zu bestimmen. Bei einer linearen Abhängigkeit der Form

$$y = y(x) = a + bx$$

lassen sich die Parameter a und b bestimmen. Ein allgemeines Verfahren, um solche Probleme zu behandeln, ist die Methode der kleinsten Quadrate. Im Falle der Anpassung einer Geraden an Daten verfährt man wie folgt.

Zu gegebenen Werten der Parameter a und b können die Residuen

$$\epsilon_i = y_i - (a + bx_i)$$

berechnet werden. Nach der Methode der kleinsten Quadrate sind optimale Schätzwerte für a und b die Werte, für die die Summe der Quadrate der Residuen

$$S(a,b) = \sum_{i=1}^{n}\epsilon_i^2 = \sum_{i=1}^{n}\left(y_i-(a+bx_i)\right)^2 \sum_{i=1}^{n}\left[y_i^2 - 2ay_i - 2bx_iy_i + a^2 + 2abx_i + b^2x_i^2\right].$$

minimal ist. Zur Bestimmung des Minimums von S bezüglich a und b werden die partiellen Ableitungen von S nach a und b gebildet:

$$
\begin{aligned}
\frac{\partial S}{\partial a} &= 2\sum_{i=1}^{n}(-y_i + a + bx_i) \\
\frac{\partial S}{\partial b} &= 2\sum_{i=1}^{n}\left(-x_i y_i + a x_i + b x_i^2\right),
\end{aligned}
$$

Die Bedingung, daß die Ableitungen am Minimum verschwinden, führt auf das lineare Gleichungssystem

$$
\begin{aligned}
an + b\sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} y_i \\
a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} x_i y_i
\end{aligned}
$$

Zur Vereinfachung werden die folgenden Summen definiert:

$$
S_x = \sum_{i=1}^{n} x_i \qquad S_{xx} = \sum_{i=1}^{n} x_i^2 \qquad S_y = \sum_{i=1}^{n} y_i \qquad S_{xy} = \sum_{i=1}^{n} x_i y_i \qquad S_{yy} = \sum_{i=1}^{n} y_i^2
$$

Mit diesen Größen lautet das lineare Gleichungssystem für a und b:

$$
\begin{aligned}
an + bS_x &= S_y \\
aS_x + bS_{xx} &= S_{xy}.
\end{aligned}
$$

Es wird gelöst durch:

$$
a = \frac{S_y S_{xx} - S_{xy} S_x}{nS_{xx} - S_x^2} \qquad b = \frac{nS_{xy} - S_y S_x}{nS_{xx} - S_x^2}
$$

Die Varianzen und Kovarianzen (die Parameter a und b sind voneinander statistisch abhängig) ergeben sich nach den Gesetzen der Fehlerfortpflanzung zu

$$
\sigma_a^2 = \sigma^2 \frac{S_{xx}}{nS_{xx} - S_x^2} \qquad \sigma_b^2 = \sigma^2 \frac{n}{nS_{xx} - S_x^2} \qquad \sigma_{ab} = \sigma^2 \frac{-S_x}{nS_{xx} - S_x^2},
$$

wenn die Standardabweichung σ der Einzeldaten y_i bekannt ist. Wenn die Standardabweichung der Einzeldaten nicht bekannt ist, kann sie durch die Formel

$$
s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - (a + bx_i))^2 = \frac{1}{n-2}\left(S_{yy} - aS_y - bS_{xy}\right)
$$

abgeschätzt werden.